**Analyzing salary prediction of professional**

**through PowerBI**

**Introduction:**

XYZ Company is a leading firm in the industry, renowned for its innovation and talented workforce. As the company expands, it faces the challenge of attracting, retaining, and fairly compensating its employees. Understanding the factors that influence salary decisions is critical to maintaining employee satisfaction and competitive advantage in the market.

**Scenario 1: Tech Industry**

XYZ Company, a rapidly growing tech firm, faces stiff competition for top talent. To stay competitive, they need to align their salary packages with industry standards. The company aims to predict salaries for software developers, data scientists, and other tech professionals by analysing factors such as experience, skills, and performance reviews. Using PowerBI, they will visualize these factors and develop predictive models for future salary adjustments. This data-driven approach will ensure competitive and fair compensation, improving employee satisfaction and retention.

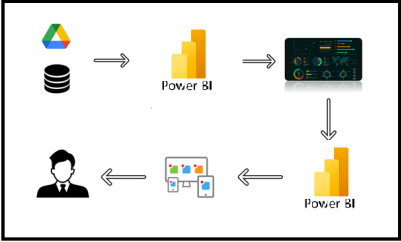
**Scenario 2: Healthcare Industry**

XYZ Company, a major player in the healthcare industry, faces challenges in compensating its diverse workforce fairly. The company aims to predict salaries for healthcare professionals by analyzing job roles, specializations, locations, and performance metrics. Using PowerBI, they will visualize salary distributions and develop predictive models for future adjustments. This approach will help address pay discrepancies and ensure competitive, equitable compensation, improving employee morale and retention.

**Scenario 3: Financial Services Industry**

XYZ Company, a leading financial services firm, needs attractive compensation packages to stay competitive. The company aims to predict salaries for financial analysts, investment bankers, and other finance professionals by analyzing educational background, certifications, experience, and performance metrics. Using PowerBI, they will visualize salary trends and create predictive models for future adjustments. This strategy will help identify and address pay gaps, ensuring competitive compensation and enhancing employee satisfaction and retention.

**Technical Architecture:**

****

**Project Flow**

To accomplish this, we have to complete all the activities listed below,

* Data Collection
  + Collect the dataset,
  + Connect Data with Power BI
* Data Preparation
  + Prepare the Data for Visualization
* Data Visualizations
  + Visualizations
* Dashboard
  + Responsive and Design of Dashboard
* Report
* Report Creation
* Performance Testing
  + Utilization of Data Filters
  + No. of Calculation fields
  + No. of Visualizations/Graphs
* Project Demonstration & Documentation
  + Record explanation Video for project end to end solution
  + Project Documentation-Step by step project development procedure

**Milestone 1: Data Collection & Extraction from Database**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

**Activity 1: Downloading the dataset**

Please use the link to download the dataset: [Click Here](https://www.kaggle.com/datasets/krishujeniya/salary-prediction-of-data-professions)

**Activity 1.1: Understand the data**

Data contains all the meta information regarding the columns described in the CSV files

**Column Description of the Dataset:**

**FIRST NAME:** The first name of the employee.

**LAST NAME:** The last name of the employee.

**SEX:** Gender of the employee.

**DOJ (Date of Joining):** The date when the employee joined the organization.

**CURRENT DATE:** The current date, used to calculate tenure.

**DESIGNATION:** The current job title of the employee.

**AGE:** The age of the employee.

**SALARY:** The current salary of the employee.

**UNIT:** The department in which the employee works.

**LEAVES USED:** The number of leave days used by the employee.

**LEAVES REMAINING:** The number of leave days remaining for the employee.

**RATINGS:** Performance ratings of the employee.

**PAST EXP (Past Experience):** The past work experience of the employee before joining the organization**.**

**Activity 2: Prepare the Data for Visualization**

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned, we can move to visualization.

3.1: Data Loading

[Click Here](https://drive.google.com/file/d/1BQrdLiTeLIKUx5jTM0fUDuXpEXjZNXI_/view?usp=sharing)

3.2 Data Cleaning

[Click Here](https://drive.google.com/file/d/1BSWUGiwocuKrJV37wDBeHoIi7jhxz3OC/view?usp=sharing)

**Milestone 2: Data Visualization**

Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

**Activity 1: World Food Production (1961-2023)**

**Activity 1.1: Employee Count**

****

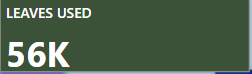
**Insight:** There are approximately 2000 employees (represented by “2K”). This metric could be relevant for organizational management, workforce planning, or evaluating company growth

**Activity 1.2: Employees Salary**

****

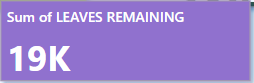
**Insight:** It significance the salary amount expressed in millions of Indian Rupees (₹143M). Such a figure could be relevant for budgeting, compensation analysis, or financial reporting within an organization.

**Activity 1.3: Leaves Used by Employees**

****

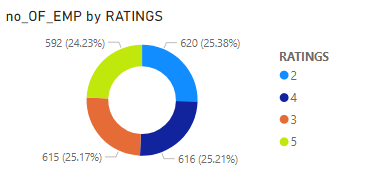
**Insight:** The “56K” could represent 56,000 leaves taken by employees—whether vacation days, sick leave, or other time off

**Activity 1.4: Remaining Leaves**

****

**Insight**: Keeping track of remaining leaves helps HR departments plan for peak seasons, holidays, and employee vacations.

**Activity 1.5: Performance Evaluation:**

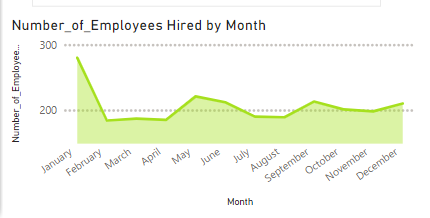
****

**Insight:** Most employees fall within the lower end of the rating scale (rating 2).

Ratings 3 and 4 are closely distributed, indicating a balanced middle ground.

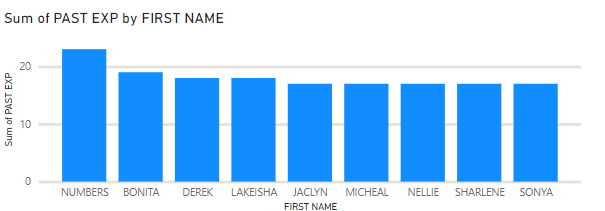
A significant number of employees have received the highest rating (5).

**Activity 1.6: Hiring Trends**

****

**Insight:** After a peak in January, there’s a decline in hiring, followed by gradual increases throughout the year. Seasonal patterns impact workforce planning

**Activity 1.7: Top 10 Experienced Employees**

****

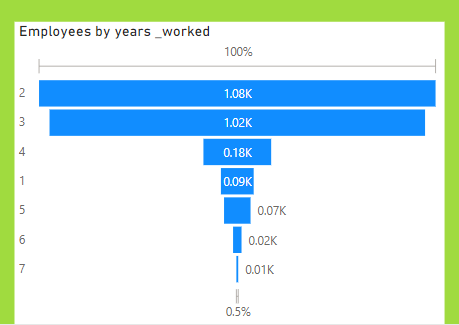
**Insight:** The column chart gives the Top 10 experienced employees in different Designation

**Activity 1.8:** Funding rounds account for 77.86% of the total investment amount, with other types such as debt funding, equity, corporate rounds, and angel rounds contributing smaller portions.

****

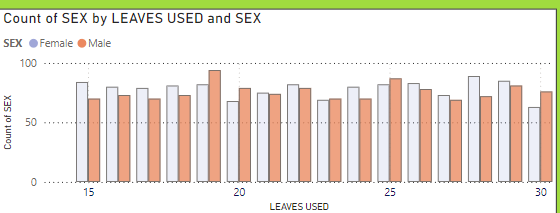
**Insight:** The slice gives an analysis of different designation such as Analyst, Associate etc.

**Activity 1.9: Experience Distribution**

****

**Insight:** Ranging from 1 to 7 years. It highlights a decreasing percentage as years of experience increase

**Activity 1.10 :**

****

**Insight:** It gives the analysis of the leaves used by differentiating the Male and Female.

**Milestone 4: Dashboard**

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key form of charts, graphs, and tables.

**Activity 1- Responsive and Design of Dashboard**

**Explanation video link:**

[**Click Here**](https://drive.google.com/file/d/1BWHv3bTiaBWr2aZBPYMz3fYJwlHjFkuz/view?usp=drive_link)

**Dashboard:**



**Insights:**

The provided PowerBI dashboard offers a comprehensive overview of employee details within a company. Here's a breakdown of the key insights:

The summary metrics show that the company employs 2,000 individuals, with 56,000 leaves used and 19,000 leaves remaining. The total salary disbursed is ₹143 million, focusing on female employees. Aaron has the highest salary of 388,000.

The employee ratings distribution is evenly spread, with 25.38% of employees rated 5, 25.21% rated 4, 25.17% rated 3, and 24.23% rated 2. This indicates a balanced performance evaluation system.

Monthly hiring trends reveal the highest hiring in January and February, peaking at nearly 300 employees in February. From March to December, the hiring rate stabilizes with slight increases in August and October, useful for planning future recruitment.

The bar graph of past experience by first name shows 'Numbers' with the highest sum of 20, followed by Bonita, Derek, Lakeisha, Jaclyn, Micheal, Nellie, Sharlene, and Sonya. This helps in project assignments and team formation based on experience.

The designation distribution includes roles like Analyst, Associate, Director, Manager, and two Senior roles, providing a clear view of the organizational structure and responsibility distribution.

In summary, this dashboard is a vital tool for HR and management, aiding in resource planning, compensation benchmarking, performance evaluation, recruitment strategies, and understanding the organizational structure.

**Milestone 5: Report**

A report is a comprehensive document that provides a detailed and structured account of data analysis, findings, and insights. It is typically used for in-depth analysis, documentation, and communication of results. Reports are suitable for a diverse audience, including decision-makers, analysts, and stakeholders who need a comprehensive understanding of the data.

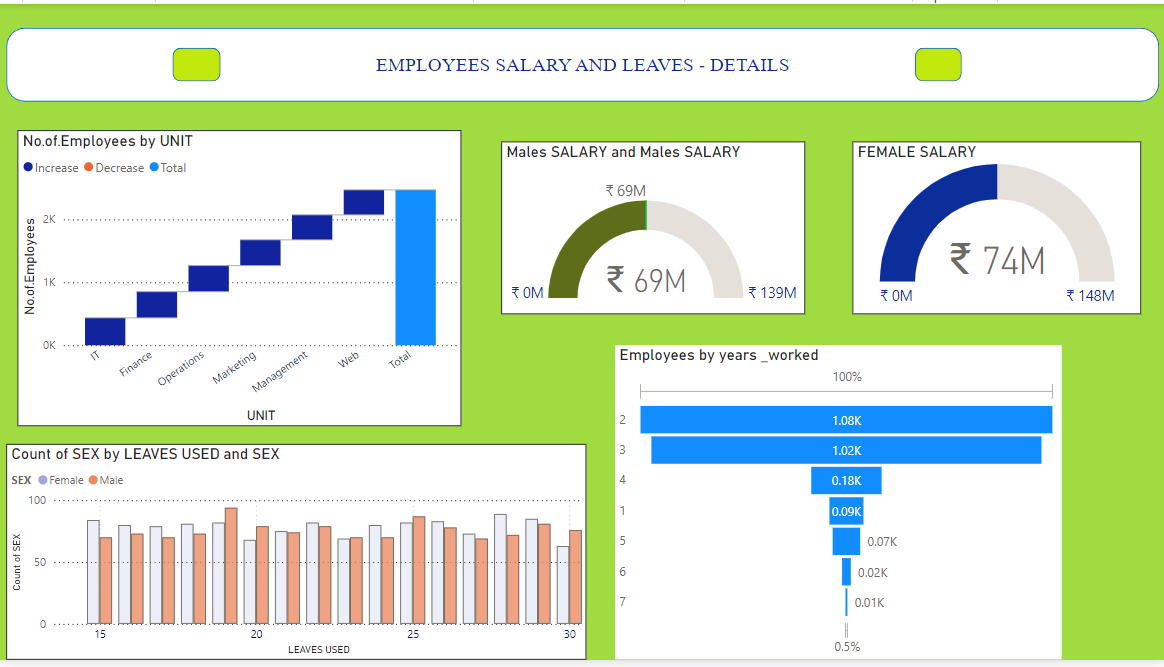
**Activity 1: Design of Report**

Designing a report in Power BI involves connecting to data sources, creating visualizations like charts and graphs, customizing their appearance and interactivity, organizing them logically on the canvas, formatting elements for consistency and clarity, and optionally creating dashboards for a summarized view. Throughout the process, it's essential to consider the audience's needs and ensure the report effectively communicates insights from the data. Finally, iterate based on feedback to continually improve the report's design and usefulness.

**Explanation video link:**

[Click Here](https://drive.google.com/file/d/1BePsKDMNwdp_6s3ZqsHSXLTApz2KBZvo/view?usp=sharing)

**Report:**



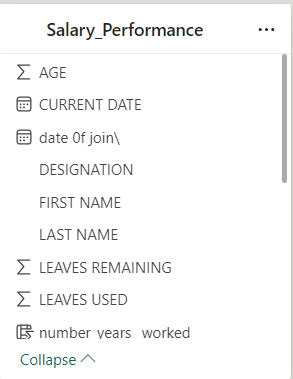
**Insights:**

The PowerBI dashboard provides a comprehensive overview of employee salary and leave details. The number of employees is distributed across various units, with the Web unit having the highest number of employees, followed by Management, Marketing, Operations, Finance, and IT. The total salary for male employees is ₹69 million, with a potential range up to ₹139 million, while female employees have a total salary of ₹74 million, extending up to ₹148 million. The leave usage comparison shows both male and female employees utilizing a similar range of leaves, indicating balanced leave usage across genders. Additionally, the distribution of employees by years worked highlights a significant portion with 1 to 2 years of experience, tapering off with fewer employees having longer tenures.

**Milestone 6: Performance Testing**

**Activity 1: Amount of Data Loaded**

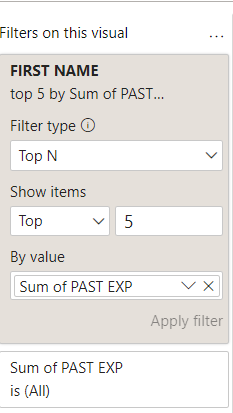
"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system.

****

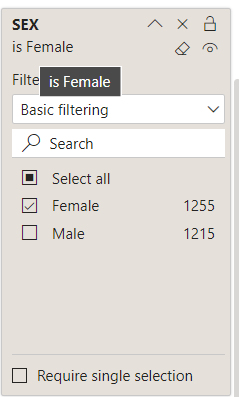
**Activity 2: Utilization of Filters**

"Utilization of Filters" involves applying filters within a system, software, or data processing pipeline to selectively extract, manipulate, or analyze data according to specific criteria or conditions.

**Activity 2.11: Selected “Top 5 Past Experience” as a Filter in column Graph**

****

**Activity 2.12: Selected “SEX - Female” as a Filter in Cards.**

****

**Activity 2.2: No of Visualizations/ Graphs**

* Employee Ratings Distribution*: Pie chart*
* Monthly Hiring Trends: *Line graph*
* Past Experience by First Name: *Bar graph*
* Designation Distribution*: Icon sets*
* Number of Employees by Unit: *Bar chart*
* Male Salary: *Gauge chart*
* Female Salary: *Gauge chart*
* Leave Usage by Gender*: Bar chart*
* Employees by Years Worked: *Bar chart*

**Milestone 7: Project Demonstration & Documentation**

Below mentioned deliverables to be submitted along with other deliverables

**Activity 1: - Record explanation Video for the project's end-to-end solution**

**Activity 2: - Project Documentation-Step by step project development procedure**